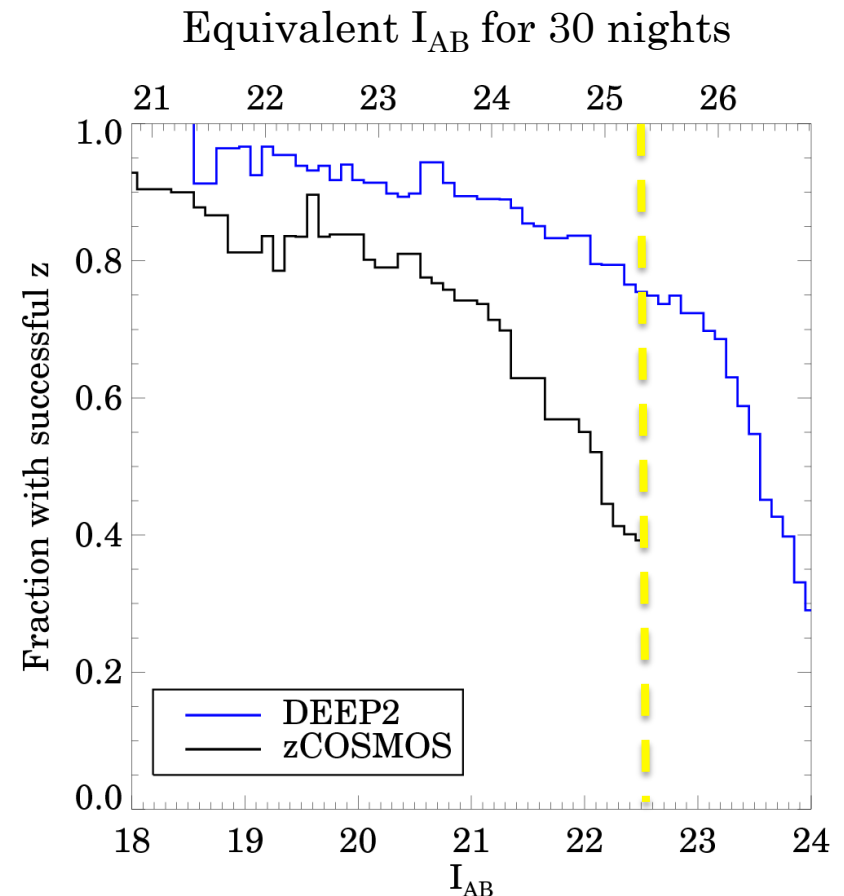


Calibrating photo-z's at LSST and WFIRST depths is limited by incompleteness in redshift surveys

- Want training set of $\sim 20\text{k}$ objects with very secure z measurements, spanning full parameter space & large volume
- As we are unlikely to achieve $>99\%$ complete calibration samples, photo- z calibration/redshift distributions would be determined via cross-correlation type methods (e.g. Newman 2008)
- For objects not spanned by training set, there's no accurate photo- z to calibrate; want $>50\%$ success at least.
- Even with instruments now being built, this will be extremely difficult from the ground at $z > 2$, degrading DE FoM



**Redshift success rates from DEEP2
(Newman et al. 2012), zCOSMOS
(Lilly et al. 2009)**

WFIRST IFU can enable Weak Lensing DE constraints from high-redshift tail

- Difficult to get secure $z > 2$ redshifts from the ground
- For LSST, DETF figure of merit (FoM) from Weak Lensing is
 - ~40% lower than ideal if only train photo- z 's to $z = 2$,
 - ~20% lower if only train to $z = 2.5$ (see Hearin et al. 2012)
- WFIRST is skewed to higher z ,
so expect similar Weak Lensing FoM degradations:
 - ~40% lower than ideal, if only train photo- z 's to $z = 2.6$,
 - ~20% lower, if only train to $z = 2.9$.
 - >70% lower FoM if only train photo- z 's to $z = 2$

IFU operations in parallel to WFIRST imaging & grism can provide large photo-z training samples

With a 3"x3" IFU

with 3" dither,

with 1.4-2ksec exposure time

expect ~10k spectra down to LSST depth.

~15k spectra down to WFIRST depth.

A 6"x6" IFU nearly doubles this.

Signal-to-noise estimates:

Typically read-noise dominated:

$$n_{read} = 76 \left(\frac{N_{exposures}}{3} \right) \left(\frac{\text{Read noise}}{3e^-} \right)^2 \left(\frac{\# \text{ spatial pixels}}{2} \right) \left(\frac{\# \text{ spec. pixels per FWHM}}{1.4} \right)$$

Tested range of scenarios:

Parameter	Optimistic	Baseline	Pessimistic
Readout noise (r_{out})	$3e^-$	$4e^-$	$4e^-$
Readout rate (f_{rout})	1/480s	1/300s	1/240s
Num. of spatial pixels (n_{pix})	2	3	4
Pixels per res. element (p_{ratio})	1.4	1.4	2.0
Total n_{read} (1440s exposure)	76 phot.	323 phot.	773 phot.

and also tried scenario where IFU goes to 2.4 μm wavelength cutoff.

WFIRST can train photo-z's at $z > 2$ with IFU

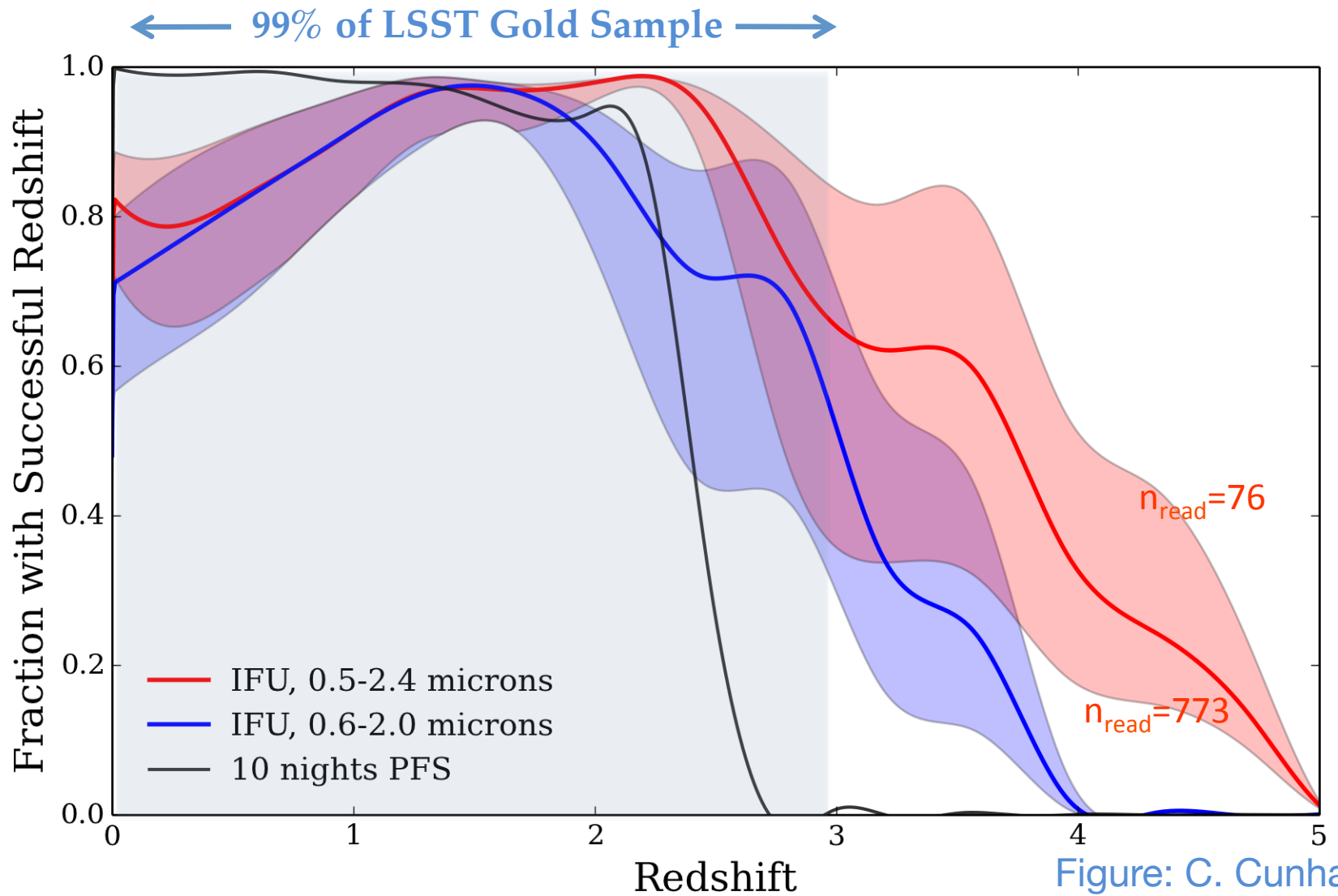
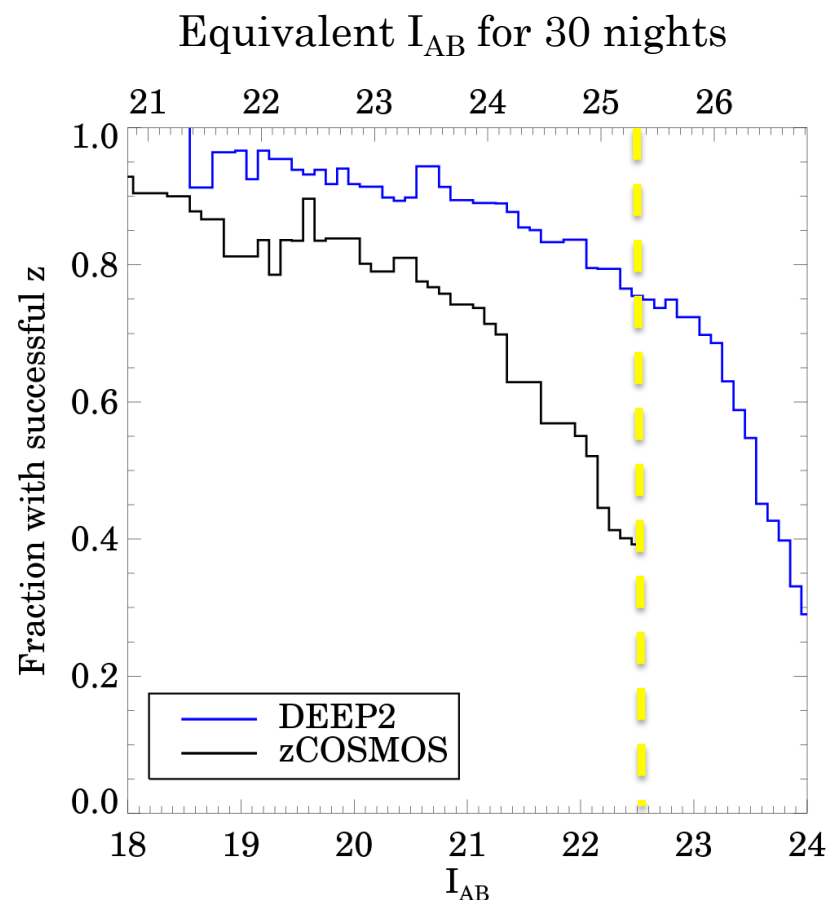


Figure: C. Cunha

Calibrating photo-z's at LSST (or WFIRST) depth is limited by incompleteness in redshift surveys

- Want training set of $\sim 20k$ objects with very secure z measurements, spanning full parameter space & large volume
- As we are unlikely to achieve $>99\%$ complete calibration samples, photo- z calibration/redshift distributions would be determined via cross-correlation type methods (e.g. Newman 2008)
- For objects not spanned by training set, there's no accurate photo- z to calibrate; want 50-75% success at least.
- Even with instruments now being built, this will be extremely difficult from the ground at $z > 2$, degrading DE FoM



Redshift success rates from DEEP2 (Newman et al. 2012), zCOSMOS (Lilly et al. 2009)

Technical Notes: Photo-z Calibration Tests

- All predictions are based on the revised COSMOS Mock Catalog (Zoubian et al., in prep.), but based on comparisons to DEEP2 and 3D-HST data we made a few changes:
 - Used velocity dispersion 65 km/sec for all objects (catalog values were unrealistic). DEEP2 95% range is 40-100 km/sec.
 - Used a fixed [OII] doublet ratio of 1:1.3 (catalog values were unrealistic)
 - For all H α -based lines (all but [OII]), catalog line fluxes were divided by 3 (the factor by which mock catalog overpredicts $z \sim 1-2$ H α EW-mass relation from Fumagilli et al. 2012); dropped Ly α (predictions are very rough at this point)
 - Consider only galaxies with star-forming spectral types
 - Early-types form a negligible fraction of the catalog at $z > 1$, and should readily yield redshifts from spectral breaks + 1.6 μ m bump for WFIRST
- Investigate completeness for emission-line redshifts only
 - Spectral break information can improve z success: predictions are conservative
- >33% improvement in FoM based on Hearin et al. 2011; estimates are based on a redshift distribution with fewer objects at $z > 2.3$ than COSMOS mock catalog

Technical Notes: PFS Simulations

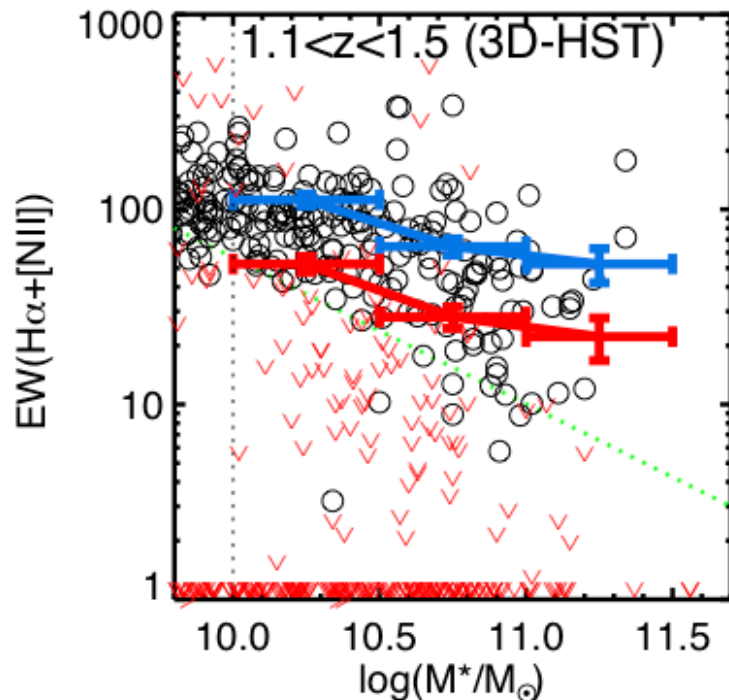
- Use resolution, pixel scales, read noise, dark current, throughput, etc. from <http://pfs.ipmu.jp/factsheet/performance.html> ; PFS-provided sky background adjusted by 1.8x to match observed DEEP2 sky flux rescaled according to telescope / instrument / fiber characteristics
- Assume all lines from COSMOS mock catalog are well-resolved (optionally including [OII] doublet); signal-to-noise is calculated for an optimal noise-weighted template-based feature detection. **The S/N estimates may be overoptimistic** (Ellis et al. claim $S/N \sim 8$ for a 5×10^{-17} erg/cm²/sec flux [OII] doublet in 15 minutes exposure; with greater collecting area & throughput, DEEP2 gets $S/N \sim 5$ in 1 hour at same flux).
- Assume six hours of integration time per night after overheads (e.g. weather).
- Use average of best-case (zenith) and worst-case (60° from zenith) atmospheric transmission.
- ~15 widely-separated PFS pointings (~30k useful spectra) are required to match 20k IFU spectra, based on sample/cosmic variance calculations using QUICKCV code of Newman & Davis (2002) . *Note:* 15x20 nights = 300 nights ~ 1 year
- Require detection of 2 emission lines at 5σ , OR 3 at 3σ significance, OR 1 at 5σ plus 3 at 2.5σ , for successful redshift determination

Technical Notes: WFIRST Simulations

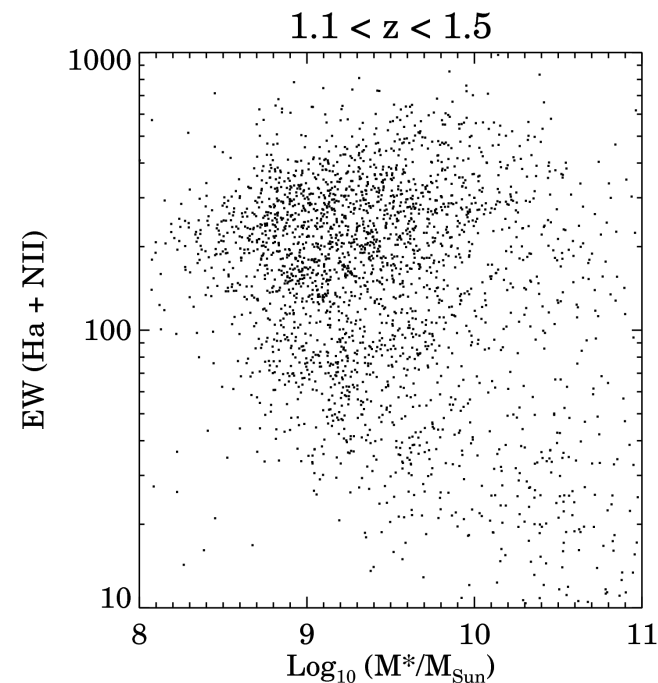
- $R=150$, 1.5 spectral pixels/resolution element, 0.5-2.4 μm wavelength coverage
- 50% throughput, 0.45'' square sky pixels, object flux distributed over 2 sky pixels
- 3 e^- RMS read noise, 0.03 e^-/s dark current, 2 readouts. Pixel scale and CCD characteristics not optimized for 2.5m.
- Assume cannot resolve [OII] or [SII] doublets; add $P\alpha/\beta/\gamma$ with Case B flux ratios vs. H α to COSMOS mock catalog line set
- Assume can always place an $i<25.3$ galaxy on IFU during imaging operations
 - For LSST “gold sample”, 40 galaxies/arcmin² = 1 per 90 arcsec²; for a 6''x6'' IFU, need to be able to move pointing center 2'' to average 1 target per pointing
- Assume 5000 deg² survey, 0.25deg² FOV; 20k imaging pointings, 2000s exposure time
- Expect 1000 5-hour and 1000 2-hour SN spectra
 - Assume can pick up another galaxy as well as host >50% of time
- Little effort made to optimize IFU parameters for photo-z calibration
- Require significant detections of 2, 3, or 4 lines with net false redshift rate, combining all criteria, of 0.1%

Flux overprediction in COSMOS mock catalogs

- Revised COSMOS mock catalogs are being used for many spectroscopic survey predictions, but current version (Zoubian et al.) appears to have significant issues.
- Catalog-derived restframe H α EW is 3x higher for star-forming galaxies than observed by 3D-HST (Fumagilli et al. 2012). Perhaps due to calibration to match Geach et al. 2010 (which overpredicts flux compared to most $z \sim 2$ samples; see their Fig. 1) ???



Fumagilli et al. 2012



Zoubian et al. (in prep.) catalog
(Figure from J. Newman)

Technical Notes: Success vs. I_{AB} : predicted vs DEEP2 rescaled

- Can compare predicted 20-night PFS redshift success to extrapolation from DEEP2, accounting for differences in collecting area and throughput
- Predicted PFS redshift success rate is somewhat higher than might expect extrapolating from DEEP2, but in the right ballpark
- $z > 1.4$ galaxies (very difficult to get with DEIMOS or VIMOS) may account for difference

